ClassTranscribe: A new tool with new educational opportunities for student crowdsourced college lecture transcription

Jia Chen Ren, Mark Hasegawa-Johnson, Lawrence Angrave

University Of Illinois at Urbana-Champaign

`jren4@illinois.edu, jhasegaw@illinois.edu, angrave@illinois.edu`

Abstract

ClassTranscribe is an open-source, web-based platform that leverages crowdsourcing to address the problem of accurate, reliable and fast transcriptions of college lectures. Completed transcriptions provide search functionality that augments existing lecture recordings and enable enhanced educational features including closed captioning.

Index Terms: crowdsourcing, human-computer interaction, lecture transcriptions

1. Introduction

Crowdsourcing research has often been explored through human marketplaces like Amazon Mechanical Turk [1] which have a core constraint that workers usually remain anonymous and transient. This constraint becomes problematic when tasks require specific domain knowledge. Attempts have been made to overcome this constraint through the use of strategies including coding theory [2] and mismatched crowdsourcing [3]. However, these strategies often come at the expense of higher cost and decreased efficiency.

Studies have shown that students produce more effective transcriptions because of the domain knowledge they possess [4]. The resulting conceptual understanding gained from the act of transcribing has been shown to be enticing [5-7] and serves as a core motivation for students to participate.

The main contribution of this work is a web-based transcription system *ClassTranscribe* that enables students to crowdsource lecture transcriptions. This system has been measured to support a transcription efficiency of $3.7\times$ and supports the parallelization of transcription tasks; no efficiency data of similar works [4,8-10] have been published. We have published the source code of ClassTranscribe online at [11] and a working demonstration at `classtranscribe.com`.

2. Design Overview

The components of ClassTranscribe are the following (Fig. 1):

- The **JobCoordinator** splits a lecture video into logical 5-10 minute transcription tasks and correspondingly assigns these tasks to participating student transcribers.

- The **FirstPass** transcription interface takes a video segment from the JobCoordinator and provides a text entry interface for transcribers to produce rough transcriptions with approximate time bounds. For more details, see Section 2.1.

- The **SecondPass** transcription interface takes a video segment along with its corresponding first pass transcription and provides a timing and text refinement interface for transcribers to produce precise and accurate final transcriptions. For more details, see Section 2.2.

- The **JobMerger** takes all finished second pass transcriptions and merges the transcriptions to produce a complete lecture transcription.
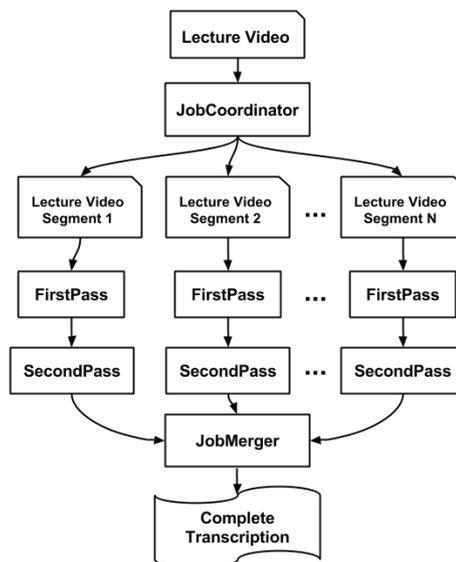

Figure 1: *The ClassTranscribe transcription process.*

2.1. FirstPass Interface

The main purpose of the first pass is to produce an accurate speech to text transcription (Fig. 2). A transcriber uses the interface by playing a lecture video at a slower playback rate ranging from $0.5\times$ to $0.75\times$ realtime speed, typing what they hear into a text box and pressing the enter key after every logical break to create a transcription segment. When a transcription segment is created, it is automatically assigned a time-range that is generated from the last transcription end time and the time when the enter key was pressed.

2.2. SecondPass Interface

The second pass interface adds precision and accuracy adjustments to the first pass transcriptions (Fig. 3). First pass transcription segments are arranged horizontally as blocks with width proportional to segment time-range length. A transcriber uses the interface by adjusting the widths of each block to match
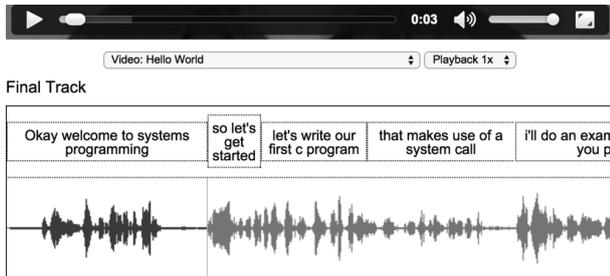
Figure 2: *FirstPass transcription interface.*

Figure 3: *SecondPass transcription interface.*

the speech timing of the lecture and correcting segment text errors.

3. Preliminary Results

The ClassTranscribe system has been deployed and tested in the University Of Illinois CS 241 Systems Programming course. The system has successfully transcribed twenty CS 241 mini videos (5-10 minutes each) and has a measured transcription efficiency of $3.7\times$, ie. every minute of lecture takes approximately 3.7 minutes to fully transcribe. Student transcribers have been observed to require a break after 25 minutes of transcribing.

The completed transcriptions have been used to develop two main educational features - closed captioning and video search. Closed captioning for lecture videos enable students with hearing disabilities and students with English as a second language to better understand and follow lectures. Video search enables students to search through all videos and directly view a video segment that covers a desired search topic. We plan to report on student interactions with these features and their effect on learning outcomes in the course.

4. Discussion

We are currently in the process of scaling ClassTrancribe to transcribe the entire semesters' CS241 lectures. This effort involves adding error checking similar to [1] using metrics collected from the transcription interfaces. We also plan to answer the following question: Does the act of transcribing a segment of lecture help reinforce a student's understanding of the concepts covered in the transcribed segment? We plan to provide ClassTranscribe to the benefit of all recorded classes at UIUC and other universities.

The speed and support for parallelism of ClassTranscribe enables a lecturer to completely transcribe a given lecture within a few hours of initial recording. Quick turnaround transcription times can enable students to review concepts quickly when reviewing for exams or completing assigned coursework using

transcription search across all lecture videos.

The ClassTranscribe platform is entirely web-based and is therefore easily accessible by students and lecturers from around the world. This coupled with the parallel nature of transcriptions tasks may enable transcription applications at a global scale.

The cost effective and scalable nature of ClassTranscribe enables the collection of a large corpus of very precise and accurate speech to text data. This data can be used for machine learning, data mining and natural language processing to open up a whole range of new possibilities for education technologies.

5. Conclusion

This work demonstrates the possibility of using students to crowdsource lecture transcriptions. The ClassTranscribe system has a measured transcription efficiency of $3.7\times$ and supports the parallelization of lecture transcription tasks. ClassTranscribe enables institutions to produce lecture transcriptions quickly, cheaply and accurately. Completed transcriptions provide search functionality that augments existing lecture recordings and enable enhanced educational features including closed captioning.

6. Acknowledgements

7. References

[1] Chia-ying Lee and James Glass, "A Transcription Task for Crowdsourcing with Automatic Quality Control," *Interspeech; 2011; Florence. ISCA; 2011. pp. 3041-5.*

[2] Aditya Vempaty, Lav R. Varshney, and Pramod K. Varshney, "Reliable Crowdsourcing for Multi-Class Labeling using Coding Theory," *IEEE J. Sel. Topics Signal Process*, vol. 8, no. 4, pp. 667–679, 2014.

[3] P. Jyothi, M. Hasegawa-Johnson, "Acquiring Speech Transcriptions Using Mismatched Crowdsourcing," *AAAI, Austin, 2015.*

[4] Raja S. Kushalnagar, Lasecki University of Rochester, Jeffrey P. Bigham. "A readability evaluation of real-time crowd captions in the classroom," *ACM Trans. Access. Comput. Article 1 2013.*

[5] Tim Causer, Justin Tonra, Valerie Wallace, "Transcription maximized; expense minimized? Crowdsourcing and editing The Collected Works of Jeremy Bentham," *Literary and Linguistic Computing, 27(2), pp. 119-137.*

[6] Galaxy Zoo, GalaxyZoo [Online] 2007. Available: http://www.galaxyzoo.org (Accessed: 25 March 2015)

[7] Sally Ellis, "A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship," *Libri. Volume 64, Issue 1, pp. 110, ISSN (Online) 1865-8423*

[8] Wald, M and Yunjia Li, "A History of Collaboration, a Future in Synote: Important Enhancements to Learning with Recorded Lectures," *DOI 10.1515/libri-2014-0001 pp. 521-525*

[9] Wald, M, "Crowdsourcing Correction of Speech Recognition Captioning Errors," *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility 2011, Article No. 22*

[10] Deshpande, R, Tuna, T., Subhlok, J. and Barker, L. "A crowdsourcing caption editor for educational videos," *Frontiers in Education Conference (FIE), 2014 IEEE, pp. 1-8*

[11] ClassTranscribe, ClassTranscribe [Online] 2015. Available: http://www.github.com/cs-education/classTranscribe (Accessed: 30 March 2015)